

10229115510000

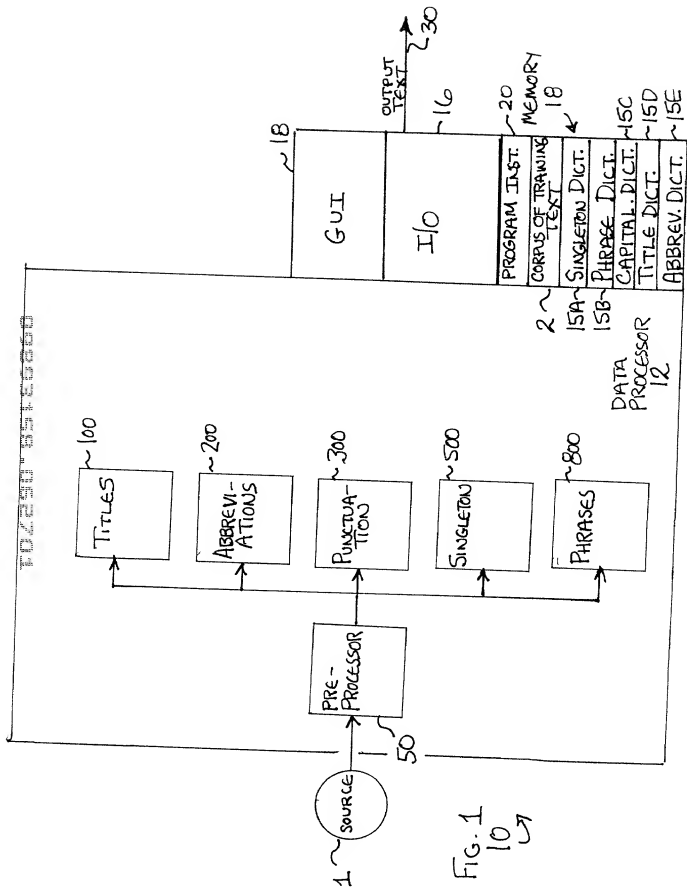


FIG. 1
10

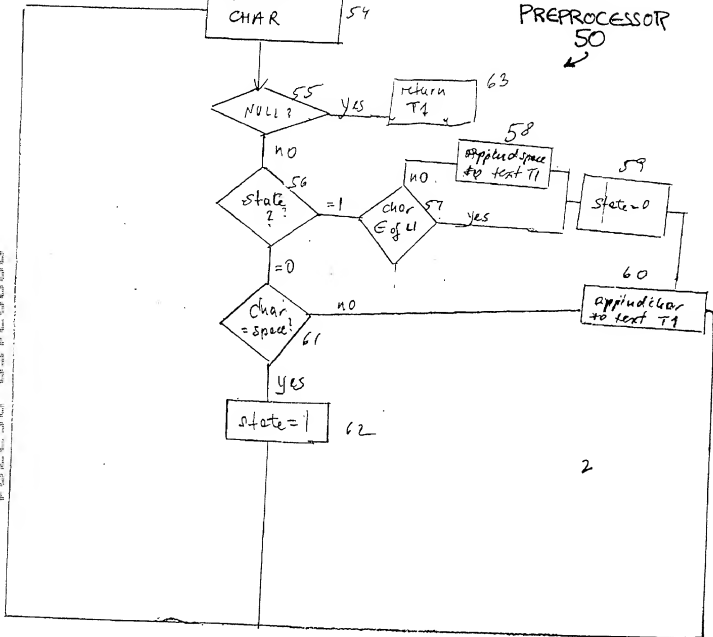
51
LIST L1

52
TEXT

53
STATE = 0

FIG. 2
PREPROCESSOR

50



0000150-062701

TEXT 52

PREPROCESSOR 50

TEXT T1 63

STATE = 1 410

GET NEXT WORD 420

NULL? 430

RETURN T2 440

APPEND WORD + SPACE TO T2 460

add SI, EI 450

PUNCTUATION PROCESSING 380

SI EI W-STRING 360

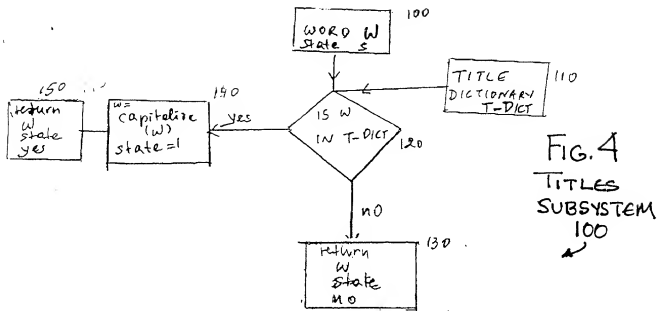
TITLE PROCESSING OF W-STRING 100

ABB-PROCESSING 200

SINGLE PROCESSING 500

FIG. 3
CAPITALIZATION
SYSTEM
10

box 600



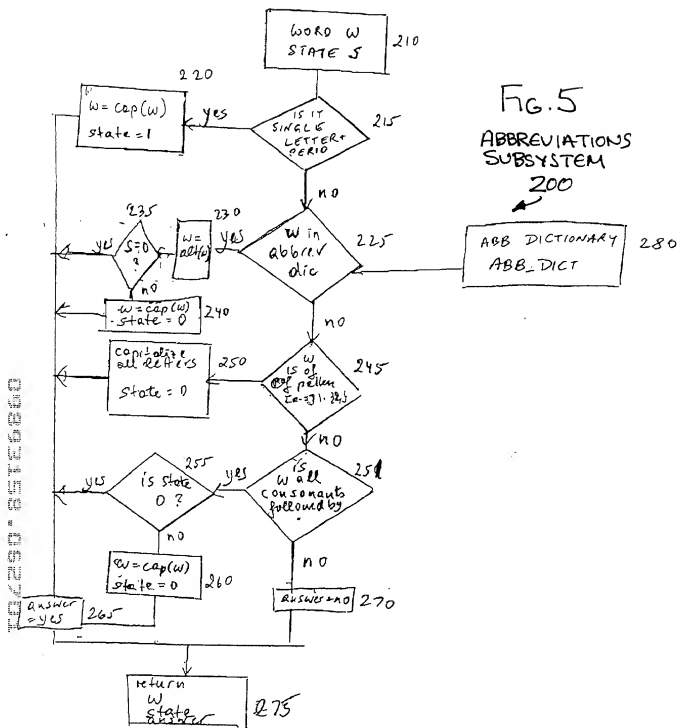
	1020	1030	1040	1050
	Term	Min Phrase Length	Max Phrase Length	Preferred Spelling
1010 {				
	⋮			

Singleton Dictionary

The diagram shows a table structure labeled 1110. It has two main columns: 1120, labeled "Phrase", and 1130, labeled "Preferred Spelling". The table contains several rows, with vertical ellipsis dots indicating more rows than shown.

	1120 Phrase	1130 Preferred Spelling
{ 1110		
	⋮	

Phrase Dictionary



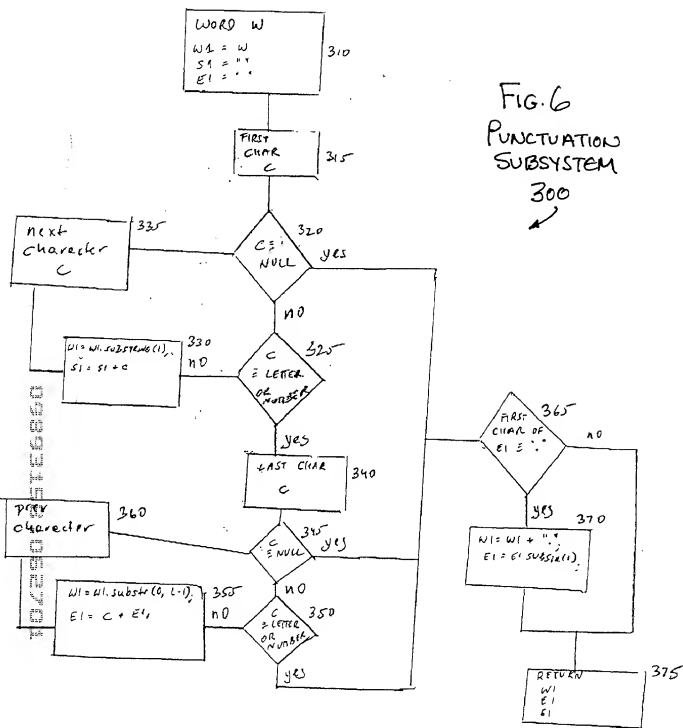


FIG. 7
SINGLETON
SUBSYSTEM

500

7

WORD W
STATE S 510

515
LAST
CHAR
= .
yes
no

520
W = W.SUBSTR(0, L-1)
S1 = ".";

525
W
ENDS WITH
"s"
yes
no

530
W = W.SUBSTR(0, L-1)
S1 = "s" + S1

535
W
in
DICT
yes
no

540
W = preferred
spelling

550
W = caplwd

540
W = alternate
spelling

555
STATE
= 1
yes
no

575
Return
W = W + S1
State

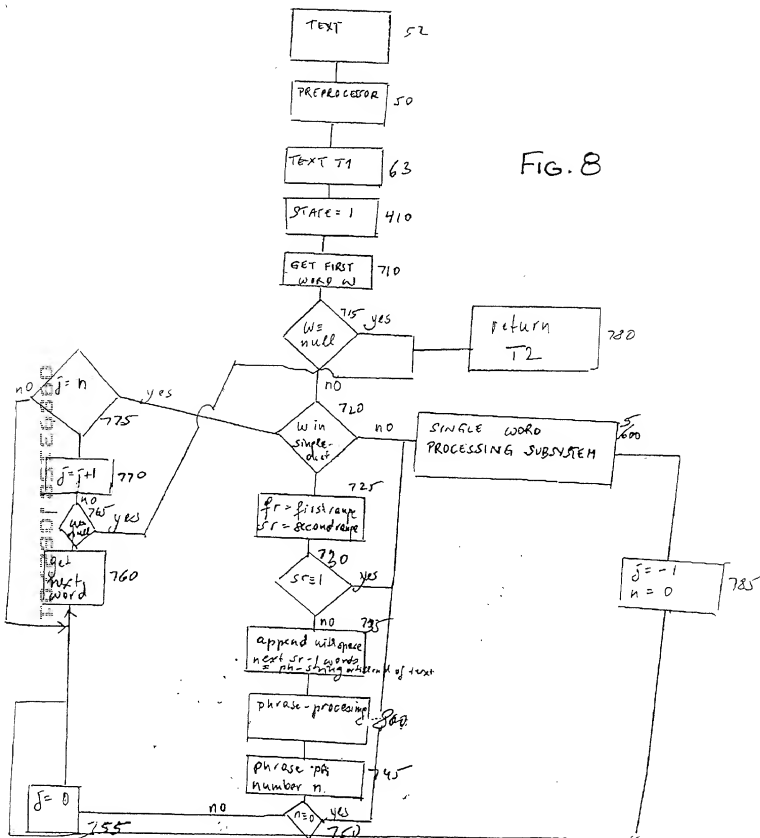
565
State = 1
yes

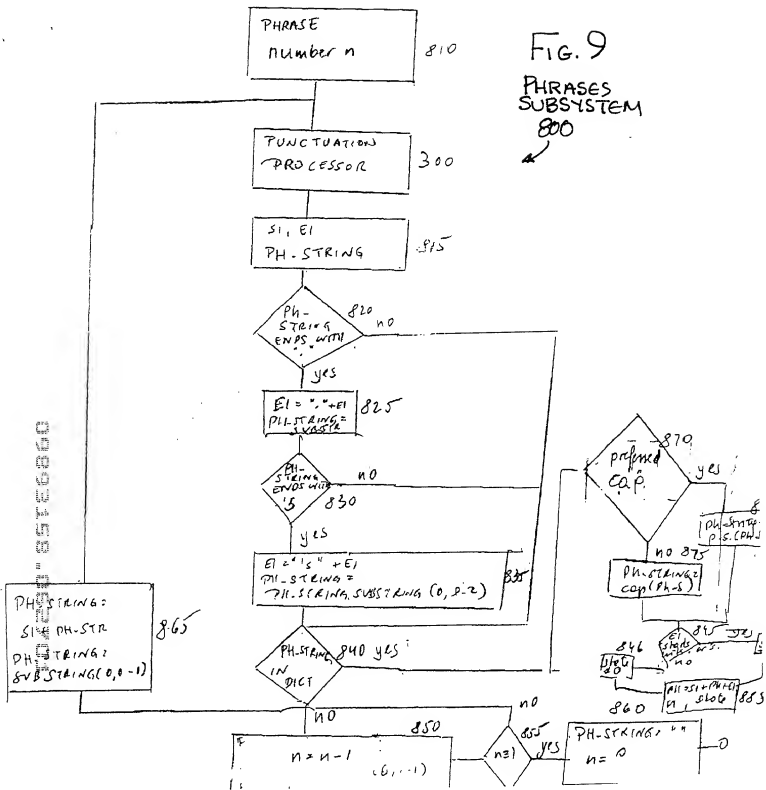
570
State = 0
no

560
S1
starts
with
"s"
yes
no

560
S1
starts
with
"s"

FIG. 8





1210 → Properly capitalized text { 2 CORPUS OF TRAINING TEXT

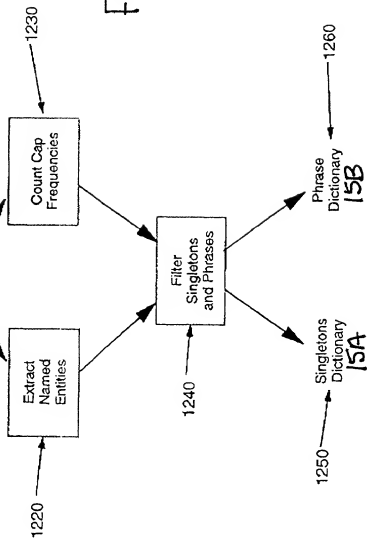


FIG. 12

Dictionary Build Process